

# Reconstruction of Correlated Sources with Energy Harvesting Constraints in Delay-constrained and Delay-tolerant Communication Scenarios

Miguel Calvo-Fullana, Javier Matamoros, and Carles Antón-Haro

**Abstract**—In this paper, we investigate the reconstruction of time-correlated sources in a point-to-point communications scenario comprising an energy-harvesting sensor and a Fusion Center (FC). Our goal is to minimize the average distortion in the reconstructed observations by using data from previously encoded sources as side information. First, we analyze a delay-constrained scenario, where the sources must be reconstructed before the next time slot. We formulate the problem in a convex optimization framework and derive the optimal transmission (i.e., power and rate allocation) policy. To solve this problem, we propose an iterative algorithm based on the subgradient method. Interestingly, the solution to the problem consists of a coupling between a two-dimensional directional water-filling algorithm (for power allocation) and a reverse water-filling algorithm (for rate allocation). Then we find a more general solution to this problem in a delay-tolerant scenario where the time horizon for source reconstruction is extended to multiple time slots. Finally, we provide some numerical results that illustrate the impact of delay and correlation in the power and rate allocation policies, and in the resulting reconstruction distortion. We also discuss the performance gap exhibited by a heuristic *online* policy derived from the optimal (offline) one.

## I. INTRODUCTION

Sensor nodes are usually powered by batteries which can be costly, difficult or even impossible to replace (e.g., when nodes are deployed in remote locations). In recent years, energy harvesting has emerged as a technology capable of overcoming (or, at least, alleviating) the limitations imposed by non-rechargeable batteries. Specifically, nodes equipped with an energy harvesting device are capable of scavenging e.g., solar, wind, thermal, kinetic energy from the environment [1] and, by doing so, extend their operational lifetime.

Energy harvesting has received considerable attention by the wireless communications and information theory communities (see [2] and references therein for an overview of current advances). For point-to-point scenarios, and under the assumption of known energy and data arrivals (*offline* optimization), the main focus has been on the derivation of optimal transmission strategies at the sensor node. In [3], the authors study the problem of minimizing the time by which all data packets are transmitted to the destination.

This work was partly sponsored by the Catalan Government under grant SGR2014-1567, the Spanish Government under projects PCIN-2013-027 (E-CROPS) and TEC2013-44591-P (INTENSIV), and the European Commission under Grant Agreement 318306 (NEWCOM#).

The authors are with the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), 08860 Castelldefels, Barcelona, Spain (e-mail: {miguel.calvo, javier.matamoros, carles.anton}@cttc.cat).

A number of authors go one step beyond and investigate the impact of *finite* energy storage capacity [4] or battery leakage [5]; generalize the analysis to fading channels [6]; or explicitly take into consideration the energy needed for data processing (in addition to data transmission) [7]. Further, other communication scenarios have been investigated such as the broadcast [8] or the multiple access [9] channels.

For uncoded transmissions, [10] investigates a number of energy-related aspects in a context of wireless sensor networks for parameter estimation. As for the coded case, in [11] the authors generalize Wyner-Ziv's source coding strategies with side information [12] to tree-structured sensor networks.

Several aspects of source and channel coding have been analyzed in *energy harvesting* scenarios. A point-to-point case was studied in [13], where rate-distortion allocation is optimized for stationary energy arrivals under data queue stability. These results were extended in [14] to the case of *finite* energy and data buffers. Besides, the multi-hop scenario was studied in [15] with *correlated* sources and *distributed* source coding. From a finite-horizon point of view, in [16] the problem of minimizing the reconstruction distortion of a Gaussian source is considered.

### A. Contribution

In this paper, we investigate the reconstruction of time-correlated sources in a point-to-point communications scenario<sup>1</sup>. As in [3], [5], [8], [16] we assume that energy arrivals are non-causally known, thus taking an *offline* optimization approach to the problem. Hence, the solution turns out to be a benchmark against which *any* online policy can be compared (we also propose one heuristic online policy here). Overall, the main contributions of this work are as follows

- We consider *time-correlated* sources. The introduction of temporal correlation in the sources is particularly relevant for video coding applications [17] since, in this case, images (i.e., sources) in consecutive frames are clearly correlated. Video source coding has been widely investigated in the literature [18]–[20]. In [18] the authors model video signals as a sequence of time-correlated (correlation given by a first-order auto-regressive process) spatially independent and identically distributed Gaussian processes (namely, frames). Such correlation model,

<sup>1</sup>Other communication scenarios such as multiple-access (MAC) channels, which are indeed relevant for wireless sensor networks, are left for future work in this area.

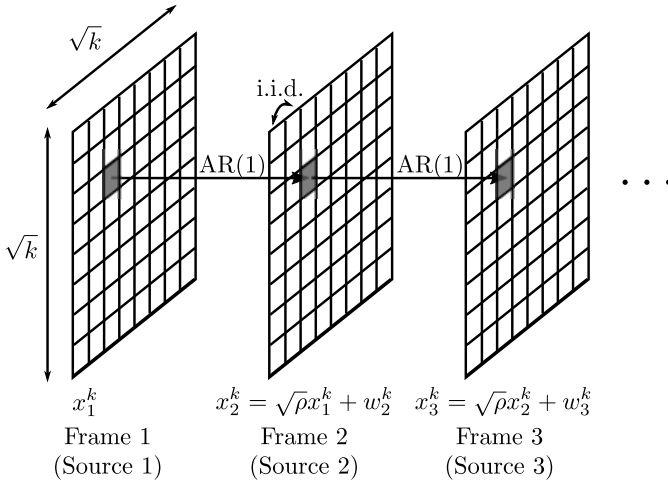


Fig. 1. Temporal and spatial correlation models in video coding.

which we adopt in this work, is illustrated in Fig. 1. Still, more general correlation models are also available [20].

- We consider *delay-tolerant* reconstruction scenarios. In other works [19], the authors analyze the impact of a *delay-tolerant* reconstruction of the correlated source. Those studies, however, were conducted in scenarios without energy harvesting. Our work goes one step beyond and incorporates energy harvesting constraints (in the sensor node) in the source coding process itself. Consequently, the closed-form expressions that we derive for the power and rate allocation policies explicitly take correlation into account. In this respect, we generalize the results of [16] to the correlated case.
- In contrast to previous works, we further leverage on *side information-aware* coding strategies for WSNs [11] to exploit correlation. First, we study the *delay-constrained* case [21] in which the source must be reconstructed at the FC before the next time slot. Then we generalize our study to the *delay-tolerant* case—where the time horizon for source reconstruction is extended to multiple time slots. For both cases, we derive the optimal transmission policy which minimizes the average reconstruction distortion at the destination. Our policy reverts to that of [16] for uncorrelated sources, and to that of [3] for the uncorrelated and delay-constrained case. In order to compute this rate and power allocation policy, we propose an iterative algorithm based on the subgradient method [22]. Interestingly, we show that the procedure encompasses the interaction between a two-directional directional water-filling and a reverse water-filling [23, Chapter 10] schemes.

Besides, we provide extensive numerical results which illustrate the impact of correlation and delay in the transmit policy and the resulting reconstruction distortion.

The remainder of this paper is organized as follows. In Section II we introduce the system model and provide details on the encoding process. In Section III, we address the distortion minimization problem for the delay-constrained case. We formulate the problem as a convex program and

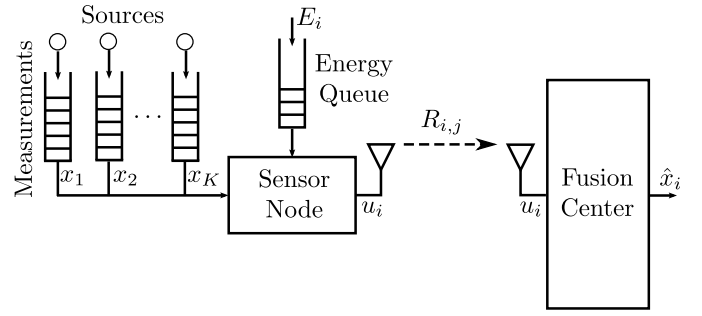


Fig. 2. System Model.

derive the optimal power and rate allocation policy. In order to compute this resulting transmission policy, we propose in Section III-C an iterative algorithm based on the subgradient method. Next, in Section IV we generalize the problem (and the solution) to the delay-tolerant case. We provide numerical results in Section V, where the effect of correlation as well as delay on the transmit policy and the resulting distortion are assessed. Finally, Section VI closes the paper by providing some concluding remarks.

## II. SYSTEM MODEL

Consider the point-to-point communications scenario depicted in Fig. 2 which comprises one energy-harvesting (EH) sensor and one Fusion Center (FC). We adopt a slotted transmission model, with  $K$  denoting the total number of time slots. The sensor measures a time-varying phenomenon of interest which, in the sequel, we model by multiple correlated and memoryless Gaussian wide-sense stationary sources (see rationale in the preceding section). Specifically, each source models the phenomenon in a given time slot. In the  $k$ -th time slot, the sensor node (i) collects a *large* number of independent and identically distributed (i.i.d.) samples from the  $k$ -th source; and (ii) encodes those measurements. The encoded data is then transmitted to the FC in  $d$  consecutive time slots.

In this work, we consider both delay-constrained ( $d = 1$ ) and delay-tolerant ( $d > 1$ ) communication scenarios. Clearly, in delay-tolerant scenarios the encoded data transmitted in a given time slot corresponds to multiple sources, as Figure 3 illustrates. Let  $R_{i,j}$  denote the average transmission rate assigned to the encoded samples of the  $j$ -th source in the  $i$ -th time slot. Necessarily, the sum-rate in the  $i$ -th time slot is upper bounded by the channel capacity<sup>2</sup>, namely,

$$\sum_{j=i-d+1}^i R_{i,j} \leq \log(1 + |h_i|^2 p_i), \quad i = 1, \dots, K, \quad (1)$$

with  $|h_i|^2$  and  $p_i$  standing for the channel gain and average transmit power in time slot  $i$ , respectively (channel noise is assumed to be Gaussian-distributed, with zero-mean and unit variance). The  $n$  i.i.d. samples collected by the sensor node from the  $i$ -th source will be denoted in the sequel by

<sup>2</sup>For the ease of notation, we let the number of channel uses to be equal to the number of samples collected in a given time slot. This number, in turn, is assumed to be large enough to satisfy Shannon's source coding theorem.

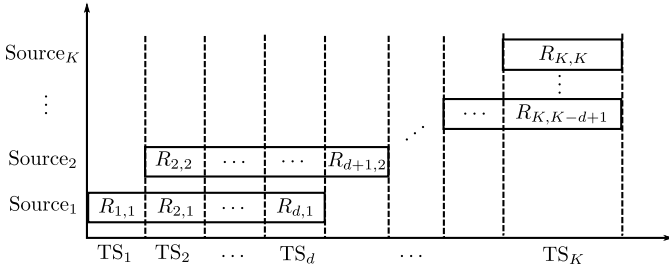


Fig. 3. Simultaneous transmission of source measurements.

$\{x_i^k\}_{k=1}^n$ . Such samples, we assume, are correlated over time slots through a first-order autoregressive process. Hence, for the  $k$ -th sample from the  $i$ -th source we have that

$$x_i^k = \sqrt{\rho} x_{i-1}^k + w_i^k, \quad \begin{matrix} k = 1, \dots, n, \\ i = 1, \dots, K, \end{matrix} \quad (2)$$

with  $\rho = \mathbb{E}[x_i^k x_{i-1}^k]$  denoting the correlation coefficient, and  $w_i^k$  standing for an i.i.d. zero-mean Gaussian random variable with variance  $\sigma_w^2 = (1 - \rho)\sigma_x^2$ .

As for the underlying energy harvesting process, we model it as a counting process [3], [4] with packet energy arrivals of  $E_i$  Joules at the beginning of time slot  $i$ . For simplicity, we assume that energy can be stored in a rechargeable battery of infinite capacity. By considering transmit power as the only energy cost, any transmission (power allocation) policy  $\{p_i\}$  at the sensor node must satisfy the following energy causality constraint:

$$T_s \sum_{j=1}^i p_j \leq \sum_{j=1}^i E_j, \quad i = 1, \dots, K, \quad (3)$$

where  $T_s$  denotes the duration of the time slot which, in the sequel, we normalize (i.e.,  $T_s = 1$ ).

**Remark 1.** There exist more sophisticated power consumption models encompassing non-ideal circuit power consumption effects [24] or the impact of processing power [7] (See also references [25] and [26] for more insight into these matters). In this work, for simplicity, we restrict ourselves to *transmit* power consumption. Nonetheless, the proposed model easily adapts to a constant circuit power consumption. Any model of the form  $T_s \sum_{j=1}^i p_j + T_s \sum_{j=1}^i P_i^c \leq \sum_{j=1}^i E_j$ , with  $P_i^c$  being the circuit power consumption at the  $i$ -th time slot, can be rewritten as  $T_s \sum_{j=1}^i p_j \leq \sum_{j=1}^i \bar{E}_j$ , where we have defined a new energy harvesting process as  $\bar{E}_i = E_i - T_s P_i^c$ . Using this new energy harvesting process, the proposed framework can be used to obtain policies adapted to a constant circuit power consumption.

Our goal is to reconstruct at the FC the sequence of measurements  $\{x_i^k\}_{k=1}^n$  of each source in up to  $d$  time slots since they were collected. Due to the continuous-valued nature of the sources and the rate constraint (1), the reconstructed measurements  $\{\hat{x}_i^k\}_{k=1}^n$  will be unavoidably subject to some distortion. Such distortion will be characterized by a Mean Squared Error (MSE) metric:

$$D_i = \frac{1}{n} \sum_{k=1}^n (x_i^k - \hat{x}_i^k)^2, \quad i = 1, \dots, K. \quad (4)$$

### A. Source Coding and Distortion

Hereinafter, we assume separability of source and channel coding at the sensor node. Hence,  $\{x_i^k\}_{k=1}^n$  can be first encoded into a length- $n$  codeword (with a sufficiently large  $n$ ) given by  $\{u_i^k\}_{k=1}^n$ . This process, as in [27], can be modeled as

$$u_i = x_i + z_i, \quad i = 1, \dots, K, \quad (5)$$

where  $z_i$  denotes i.i.d. zero-mean Gaussian random noise of variance  $\sigma_{z_i}^2$ , which plays the role of encoding noise (the sample index has been omitted here for brevity). We know that, in order to decode the received data, the FC will exploit the available side information (i.e., all the preceding  $u_i$ ). Hence, the sum of the (average) encoding rates per sample for the  $i$ -th source over the  $d$  consecutive time slots must satisfy [23]

$$\sum_{j=i}^{i+d-1} R_{j,i} \geq I(x_i; u_i | u_1, \dots, u_{i-1}), \quad i = 1, \dots, K, \quad (6)$$

where  $I(\cdot; \cdot | \cdot)$  stands for the conditional mutual information. From (5), this last expression can be rewritten as

$$\begin{aligned} I(x_i; u_i | u_1, \dots, u_{i-1}) &= H(u_i | u_1, \dots, u_{i-1}) - \\ &\quad H(u_i | u_1, \dots, u_{i-1}, x_i) \\ &= \log \left( 1 + \frac{\sigma_{x_i | u_1, \dots, u_{i-1}}^2}{\sigma_{z_i}^2} \right), \end{aligned} \quad (7)$$

with  $H(\cdot | \cdot)$  standing for the conditional entropy and  $\sigma_{x_i | u_1, \dots, u_{i-1}}^2$  for the conditional variance of the  $i$ -th observation given all the previous data available at the FC. Hence, by taking equality in (6), the variance of the encoding noise reads

$$\sigma_{z_i}^2 = \frac{\sigma_{x_i | u_1, \dots, u_{i-1}}^2}{\sum_{j=i}^{i+d-1} R_{j,i} - 1}. \quad (8)$$

In each time slot, the FC produces an optimal Minimum Mean Squared Error (MMSE) estimate of the observations which, as discussed earlier, exploits all the preceding  $u_i$ , namely

$$\hat{x}_i = \mathbb{E}[x_i | u_1, \dots, u_i], \quad i = 1, \dots, K. \quad (9)$$

The distortion in the reconstruction of  $x_i$  thus reads:

$$D_i = \sigma_{x_i | u_1, \dots, u_i}^2, \quad (10)$$

which, in turn, can be expressed as<sup>3</sup> (see Appendix A, for a detailed derivation)

$$\begin{aligned} D_i &= \sigma_x^2 \left( (1 - \rho) \sum_{j=2}^i \rho^{i-j} e^{-\sum_{k=j}^i \sum_{l=k}^{k+d-1} R_{l,k}} \right. \\ &\quad \left. + \rho^{i-1} e^{-\sum_{k=1}^i \sum_{l=k}^{k+d-1} R_{l,k}} \right). \end{aligned} \quad (11)$$

<sup>3</sup>With some abuse of notation, in the summation interval we write  $k+d-1$ . Still, we restrict such summations to the valid range of timeslot values, namely,  $\max\{k+d-1, K\}$ .

### III. MINIMIZATION OF THE AVERAGE DISTORTION: DELAY-CONSTRAINED SCENARIO

Here, samples must be encoded, transmitted and reconstructed before the next time slot starts (i.e.,  $d = 1$ ). The particularization of the channel capacity constraint (1) thus reads

$$R_i \leq \log(1 + |h_i|^2 p_i), \quad (12)$$

where  $R_i$  stands for the transmission rate which is assigned to the  $i$ -th source in the  $i$ -th time slot *only* (i.e., no summation of rates over subsequent time slots). Likewise, the rate-distortion constraint (6) can be particularized to

$$R_i \geq I(x_i; u_i | u_1, \dots, u_{i-1}), \quad (13)$$

From all the above, the reconstruction distortion in (11) simplifies to

$$D_i = \sigma_x^2 \left( (1 - \rho) \sum_{j=2}^i \rho^{i-j} e^{-\sum_{k=j}^i R_k} + \rho^{i-1} e^{-\sum_{k=1}^i R_k} \right). \quad (14)$$

Our goal is to find the optimal power  $\{p_i\}$  and rate  $\{R_i\}$  allocation that minimize the average distortion given by (14) subject to the energy causality constraint of (3) and the capacity constraint of (1). Unfortunately, due to the coupling (over time slots) of the rates in the exponential terms of (14), this optimization problem cannot be solved analytically. To circumvent this, we define the *cumulative rates*  $r_{ij}$  as  $r_{ij} \triangleq \sum_{k=j}^i R_k$ , for  $i = 1, \dots, K, j = 1, \dots, i$ . By doing so, the optimization problem can be posed as:

$$\min_{\substack{\{p_i\}, \\ \{R_i\}, \\ \{r_{ij}\}}} \frac{\sigma_x^2}{K} \sum_{i=1}^K \left( (1 - \rho) \sum_{j=2}^i \rho^{i-j} e^{-r_{ij}} + \rho^{i-1} e^{-r_{i1}} \right) \quad (15a)$$

$$\text{s.t. } r_{ij} = \sum_{k=j}^i R_k, \quad i = 1, \dots, K, j = 1, \dots, i \quad (15b)$$

$$R_i \leq \log(1 + |h_i|^2 p_i), \quad i = 1, \dots, K, \quad (15c)$$

$$\sum_{j=1}^i p_j \leq \sum_{j=1}^i E_j, \quad i = 1, \dots, K, \quad (15d)$$

$$-p_i \leq 0, \quad i = 1, \dots, K, \quad (15e)$$

$$-R_i \leq 0, \quad i = 1, \dots, K, \quad (15f)$$

$$-r_{ij} \leq 0, \quad i = 1, \dots, K, j = 1, \dots, i \quad (15g)$$

where the optimization is with respect to variables  $\{p_i\}$ ,  $\{R_i\}$  and, also,  $\{r_{ij}\}$  (this follows from the introduction of the additional constraint (15b) associated to the definition of cumulative rates). Since the objective function (15a) is convex and the constraints (15b)-(15g) define a convex feasible set, the optimization problem (15) is convex and, thus, has a global solution [28]. By satisfying the Karush-Kuhn-Tucker (KKT) conditions, we identify the necessary and sufficient conditions

for optimality. The Lagrangian of (15) reads

$$\begin{aligned} \mathcal{L} = & \frac{\sigma_x^2}{K} \sum_{i=1}^K \left( (1 - \rho) \sum_{j=2}^i \rho^{i-j} e^{-r_{ij}} + \rho^{i-1} e^{-r_{i1}} \right) \\ & + \sum_{i=1}^K \sum_{j=1}^i \lambda_{ij} \left( r_{ij} - \sum_{k=j}^i R_k \right) \\ & + \sum_{i=1}^K \mu_i (R_i - \log(1 + |h_i|^2 p_i)) \\ & + \sum_{i=1}^K \beta_i \left( \sum_{j=1}^i p_j - \sum_{j=1}^i E_j \right) \\ & - \sum_{i=1}^K \eta_i p_i - \sum_{i=1}^K \phi_i R_i - \sum_{i=1}^K \sum_{j=1}^i \delta_{ij} r_{ij}, \end{aligned} \quad (16)$$

where  $\{\mu_i\} \geq 0$ ,  $\{\beta_i\} \geq 0$ ,  $\{\eta_i\} \geq 0$ ,  $\{\phi_i\} \geq 0$ ,  $\{\delta_{ij}\} \geq 0$  and  $\{\lambda_{ij}\}$  stand for the corresponding Lagrange multipliers. The additional complementary slackness conditions are given by

$$\mu_i (R_i - \log(1 + |h_i|^2 p_i)) = 0, \quad \forall i, \quad (17)$$

$$\beta_i \left( \sum_{j=1}^i p_j - \sum_{j=1}^i E_j \right) = 0, \quad \forall i, \quad (18)$$

$$\eta_i p_i = 0, \quad \forall i, \quad (19)$$

$$\phi_i R_i = 0, \quad \forall i, \quad (20)$$

$$\delta_{ij} r_{ij} = 0, \quad \forall i, j. \quad (21)$$

Finally, by taking the derivative of the Lagrangian with respect to  $p_i$ ,  $R_i$ ,  $r_{ij}$  and letting them be equal to zero we the set of stationarity conditions follow, namely,

$$\frac{\partial \mathcal{L}}{\partial p_i} = -\frac{\mu_i |h_i|^2}{1 + |h_i|^2 p_i} + \sum_{j=i}^K \beta_j - \eta_i = 0, \quad (22)$$

$$\frac{\partial \mathcal{L}}{\partial R_i} = -\sum_{k=i}^K \sum_{j=1}^i \lambda_{kj} + \mu_i - \phi_i = 0, \quad (23)$$

$$\frac{\partial \mathcal{L}}{\partial r_{ij}} = \begin{cases} -\frac{\sigma_x^2}{K} \rho^{i-j} e^{-r_{ij}} + \lambda_{ij} - \delta_{ij} = 0, & \text{if } j = 1, \\ -\frac{\sigma_x^2}{K} (1 - \rho) \rho^{i-j} e^{-r_{ij}} + \lambda_{ij} - \delta_{ij} = 0, & \text{if } j \neq 1. \end{cases} \quad (24)$$

#### A. Optimal Power Allocation

From the stationarity conditions on  $p_i$  (22) and  $R_i$  (23), and the slackness conditions of (21), the optimal power allocation follows:

$$p_i^* = \left[ \frac{\sum_{k=i}^K \sum_{j=1}^i \lambda_{kj}}{\sum_{j=i}^K \beta_j} - \frac{1}{|h_i|^2} \right]^+, \quad i = 1, \dots, K, \quad (25)$$

where  $[\cdot]^+ = \max\{\cdot, 0\}$ . This solution can be interpreted as a two-dimensional directional waterfilling, as shown in Fig. 4(a). For each time slot  $i$ , we have a rectangle of

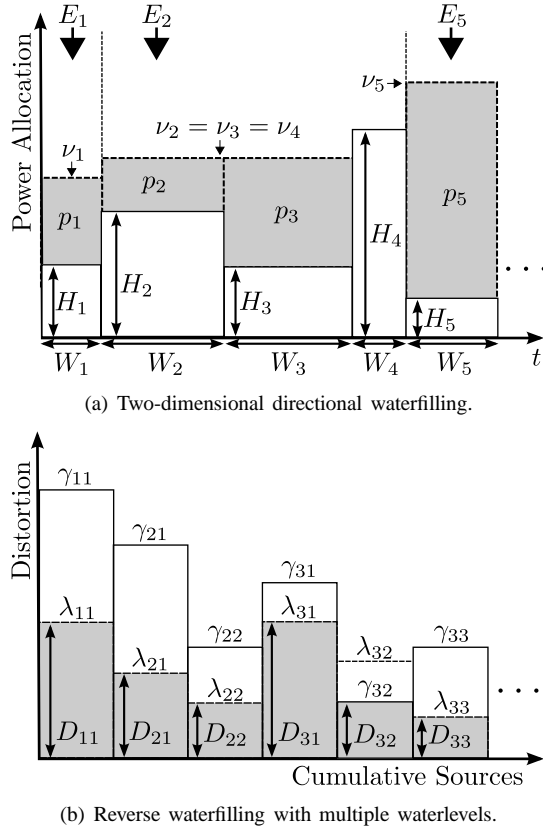


Fig. 4. Optimal power and cumulative rate allocation.

solid material of width  $W_i \triangleq \sum_{k=i}^K \sum_{j=1}^i \lambda_{kj}$  and height  $H_i \triangleq 1 / (|h_i|^2 \sum_{k=i}^K \sum_{j=1}^i \lambda_{kj})$ . Right-permeable taps are placed in time slots with energy arrivals. Water is consequently poured up to a waterlevel  $\nu_i \triangleq 1 / \sum_{j=i}^K \beta_j$ . The resulting power allocation corresponds to the area of water above the solid rectangle.

### B. Optimal Rate Allocation

Next, by solving (24) for  $r_{ij}$ , and taking into account the corresponding slackness conditions, the optimal cumulative rate allocation can be written as

$$r_{ij}^* = \begin{cases} \left[ \log \left( \frac{\frac{1}{K} \sigma_x^2 \rho^{i-j}}{\lambda_{ij}} \right) \right]^+, & \text{if } j = 1, \\ \left[ \log \left( \frac{\frac{1}{K} \sigma_x^2 (1 - \rho) \rho^{i-j}}{\lambda_{ij}} \right) \right]^+, & \text{if } j \neq 1. \end{cases} \quad (26)$$

From this last expression, it becomes apparent that, necessarily,  $\{\lambda_{ij}\} > 0$ . Hence, from (23), we have that  $\{\mu_i\} > 0$  too. This implies that constraint (15c) is satisfied with equality. Moreover, expression (26) can be readily interpreted in terms of a *reverse* water-filling solution for the reconstruction of parallel Gaussian sources [23, Chapter 10]. To see that, we define

$$\gamma_{ij} = \begin{cases} \frac{1}{K} \sigma_x^2 \rho^{i-j}, & \text{if } j = 1, \\ \frac{1}{K} \sigma_x^2 (1 - \rho) \rho^{i-j}, & \text{if } j \neq 1, \end{cases} \quad (27)$$

and

$$D_{ij} = \begin{cases} \lambda_{ij}, & \text{if } \lambda_{ij} < \gamma_{ij}, \\ \gamma_{ij}, & \text{if } \lambda_{ij} \geq \gamma_{ij}. \end{cases} \quad (28)$$

**Algorithm 1** Optimal power and rate allocation for the delay-constrained case.

- 1: **Initialize:**  $\{\lambda_{ij}^{(t)}\} := 0$ .
- 2: **Step 1:** For all  $i$ , allocate power.
- 3:  $p_i^{(t+1)} := \left[ \frac{\sum_{k=i}^K \sum_{j=1}^i \lambda_{kj}^{(t)}}{\sum_{j=i}^K \beta_j} - \frac{1}{|h_i|^2} \right]^+$
- 4: **Step 2:** For all  $i, j$ , cumulative rate allocation.
- 5:  $r_{ij}^{(t+1)} := \left[ \log \left( \frac{\gamma_{ij}}{\lambda_{ij}^{(t)}} \right) \right]^+$
- 6: **Step 3:** For all  $i, j$ , update multiplier.
- 7:  $\lambda_{ij}^{(t+1)} := \left[ \lambda_{ij}^{(t)} + \alpha \left( r_{ij}^{(t+1)} - \sum_{k=j}^i \log \left( 1 + |h_k|^2 p_k^{(t+1)} \right) \right) \right]^+$
- 8: **Step 4:** Go to Step 1 until stopping criteria is met.

Bearing the above in mind, equation (26) can be rewritten as

$$r_{ij}^* = \left[ \log \left( \frac{\gamma_{ij}}{D_{ij}} \right) \right]^+. \quad (29)$$

As Figure 4(b) illustrates, this solution mimics that of a rate-distortion allocation problem for parallel Gaussian sources. However, here the allocated rates  $r_{ij}^*$  (and sources) are *cumulative* rather than *individual*; and the reverse water level given by  $\lambda_{ij}$  is not constant. Besides, the numerator in the argument of (29) does not only depend on the variance of the sources  $\sigma_x^2$  but also on the correlation coefficient  $\rho$ , as (27) evidences.

Finally, by replacing (29) in (14), the optimal distortion for the reconstruction of the  $i$ -th source reads

$$D_i^* = \sum_{j=1}^i D_{ij}. \quad (30)$$

that is, it can be computed as the sum of the distortions associated to the corresponding cumulative rates.

### C. Optimization Algorithm

As discussed in the previous section, the optimal power (25) and cumulative rate (26) allocation are coupled<sup>4</sup> by the Lagrange multipliers  $\lambda_{ij}$ . Further, one can easily prove that problem (15) satisfies Slater's condition, and therefore, strong duality holds [28]. Since in these conditions the duality gap is zero, we propose to solve the corresponding *dual* problem in order to determine the *primal* solution (power and rates) in which we are interested. To that aim, we resort to the subgradient method [22] on which basis the solution to the dual problem,  $\{\lambda_{ij}\}$ , can be iteratively found (convergence can be guaranteed under some mild conditions). Specifically, in the  $t$ -th iteration, the Lagrange multipliers are updated as

<sup>4</sup>When dealing with hybrid power supplies (power grid and energy harvesting), two-stage waterfilling structures have also been identified in [29], [30].

$$\left\{ \begin{array}{l} r_{ii} = \sum_{j=i}^{i+d-1} R_{j,i}, i = 1, \dots, K \\ \sum_{j=i-d+1}^i R_{i,j} \leq \log(1 + |h_i|^2 p_i), i = 1, \dots, K \\ R_{i,j} \geq 0, i = 1, \dots, K, j = i-d+1, \dots, i \end{array} \right\} \equiv \left\{ \begin{array}{l} r_{ij} = \sum_{k=j}^i r_{kk}, i = 1, \dots, K, j = 1, \dots, i-1 \\ r_{ij} \leq \sum_{k=j}^{d+i-1} \log(1 + |h_k|^2 p_k), i = 1, \dots, K, j = 1, \dots, i \\ r_{ij} \geq 0, i = 1, \dots, K, j = 1, \dots, i \end{array} \right\} \quad (30)$$

follows<sup>5</sup>:

$$\lambda_{ij}^{(t+1)} := \left[ \lambda_{ij}^{(t)} + \alpha \left( r_{ij}^{(t+1)} - \sum_{k=j}^i \log(1 + |h_k|^2 p_k^{(t+1)}) \right) \right], \quad (31)$$

with  $\alpha$  standing for the corresponding step size. In Algorithm 1, we summarize the proposed procedure to solve the power and cumulative rate allocation problem.

The algorithm corresponds to a subgradient ascent on the dual function. Hence, it has a convergence rate of the order of  $\mathcal{O}(1/\sqrt{t})$  [31, Chapter 8.2]. Moreover, for a node, a single iteration of the algorithm will be of the order of  $\mathcal{O}(K \log K)$ , as it is a form of waterfilling. Namely, sorting takes  $\mathcal{O}(K \log K)$  operations, while each waterfilling operation takes  $\mathcal{O}(K)$  operations and there are at most  $\mathcal{O}(\log K)$  waterfillings to be done, since we can compute a binary search between the water bins.

#### IV. MINIMIZATION OF THE AVERAGE DISTORTION: DELAY-TOLERANT SCENARIO

Here, we address the more general case in which data is allowed to be transmitted and reconstructed within  $d > 1$  time slots after samples are collected and encoded. Again, to render the problem solvable, we define the *cumulative rates* as

$$r_{ij} \triangleq \sum_{k=j}^i \sum_{l=k}^{k+d-1} R_{l,k} \quad \text{for } i = 1, \dots, K, j = 1, \dots, i. \quad (32)$$

In delay-tolerant scenarios, each time slot conveys data from up to  $d$  different sources (see Fig. 3). Hence, the number of unknowns ( $\{R_{i,j}\}$ ),  $Kd - d(d-1)/2$  in total, exceeds the number of equations given by the capacity constraints (1),  $K$  in total. Consequently, the system of equations becomes underdetermined. This means that, even if a unique solution exists when optimizing on the *cumulative rates*  $r_{ij}$  (as we discuss next), there exist *multiple* solutions for the *individual rates*. Thus, we propose to solve the optimization problem in terms of cumulative rates (only), and then define some criteria to select one solution in terms of individual rates (this will be further elaborated in Section V ahead).

To start with, we need to rewrite not only (i) the objective function given by (11); but, also, (ii) the set of constraints, in terms of cumulative rates. The latter can be accomplished by expressing the cumulative rates in the following recursive

form:

$$r_{ij} \triangleq \begin{cases} \sum_{k=j}^i r_{kk}, & \text{if } j \neq i \\ \sum_{k=j}^{j+d-1} R_{k,j}, & \text{if } j = i \end{cases} \quad (33)$$

for  $i = 1, \dots, K, j = 1, \dots, i$ , and then resorting to Fourier-Motzkin elimination [32]. Further, we prove that the system obtained by Fourier-Motzkin elimination is equivalent.

**Proposition 2.** *The systems of inequalities in (30) are equivalent when solving optimization problem (34). That is, the set of variables  $\{p_i\}, \{R_{i,j}\}$  and  $\{r_{ij}\}$  satisfy the constraints on the left hand side of (30) if and only if they satisfy the constraints on the right hand side of (30).*

*Proof:* See Appendix B. ■

Finally, in order to pose the optimization problem, it suffices to include the corresponding *energy harvesting* constraints of (3) too, namely

$$\min_{\{p_i\}, \{r_{ij}\}} \frac{\sigma_x^2}{K} \sum_{i=1}^K \left( (1 - \rho) \sum_{j=2}^i \rho^{i-j} e^{-r_{ij}} + \rho^{i-1} e^{-r_{i1}} \right) \quad (34a)$$

$$\text{s.t. } r_{ij} = \sum_{k=j}^i r_{kk}, i = 1, \dots, K, j = 1, \dots, i-1 \quad (34b)$$

$$r_{ij} \leq \sum_{k=j}^{d+i-1} \log(1 + |h_k|^2 p_k), \quad i = 1, \dots, K, j = 1, \dots, i \quad (34c)$$

$$\sum_{j=1}^i p_j \leq \sum_{j=1}^i E_j, \quad i = 1, \dots, K, \quad (34d)$$

$$-p_i \leq 0, \quad i = 1, \dots, K, \quad (34e)$$

$$-r_{ij} \leq 0, \quad i = 1, \dots, K, j = 1, \dots, i \quad (34f)$$

where, clearly, the optimization is now with respect to variables  $\{p_i\}$  and  $\{r_{ij}\}$ . Differently from Section III, constraint (34b) guarantees, on the one hand, that the cumulative rates satisfy definition (33). On the other, constraint (34c) enforces the cumulative rates to satisfy the per time slot channel capacity constraint.

The optimization problem (34) is convex and can be solved

<sup>5</sup>Here, we use extended-value definitions for all functions [28], thus taking  $+\infty$  values outside their respective domain.

in closed-form by (i) computing the Lagrangian function:

$$\begin{aligned} \mathcal{L} = & \frac{\sigma_x^2}{K} \sum_{i=1}^K \left( (1-\rho) \sum_{j=2}^i \rho^{i-j} e^{-r_{ij}} + \rho^{i-1} e^{-r_{i1}} \right) \\ & + \sum_{i=1}^K \sum_{j=1}^i \mu_{ij} \left( r_{ij} - \sum_{k=j}^i r_{kk} \right) \\ & + \sum_{i=1}^K \sum_{j=1}^i \lambda_{ij} \left( r_{ij} - \sum_{k=j}^{d+i-1} \log(1 + |h_i|^2 p_i) \right) \\ & + \sum_{i=1}^K \beta_i \left( \sum_{j=1}^i p_j - \sum_{j=1}^i E_j \right) \\ & - \sum_{i=1}^K \eta_i p_i - \sum_{i=1}^K \sum_{j=1}^i \delta_{ij} r_{ij}. \end{aligned} \quad (35)$$

with  $\{\lambda_{ij}\} \geq 0$ ,  $\{\beta_i\} \geq 0$ ,  $\{\eta_i\} \geq 0$ ,  $\{\delta_i\} \geq 0$ ; and  $\{\mu_{ij}\}$  standing for the corresponding Lagrangian multipliers; and (ii) satisfying the Karush-Kuhn-Tucker (KKT) conditions that follow from the Lagrangian. Along the lines of Section III, the optimal power allocation reads,

$$p_i^* = \left[ \frac{\sum_{k=i-d+1}^K \sum_{l=1}^k \lambda_{kl}}{\sum_{j=i}^K \beta_j} - \frac{1}{|h_i|^2} \right]^+, \quad i = 1, \dots, K. \quad (36)$$

The optimal power allocation for the (more general) delay-tolerant scenario admits again a two-dimensional directional waterfilling interpretation. Differently from the delay-constrained scenario, the width and height of the solid rectangle, namely  $W_i \triangleq \sum_{k=i-d+1}^K \sum_{l=1}^k \lambda_{kl}$  and  $H_i \triangleq 1/(|h_i|^2 \sum_{k=i-d+1}^K \sum_{l=1}^k \lambda_{kl})$ , have an explicit dependence on  $d$ , the maximum latency<sup>6</sup>.

Along the lines of the preceding section, the optimal cumulative rates for the delay-constrained scenario follow:

$$r_{ij}^* = \left[ \log \left( \frac{\gamma_{ij}}{\lambda_{ij} + \bar{\mu}_{ij}} \right) \right]^+. \quad (37)$$

where we have defined

$$\bar{\mu}_{ij} = \begin{cases} -\sum_{k=i}^K \sum_{\substack{l=1 \\ l \neq k}}^i \mu_{kl}, & \text{if } i = j, \\ \mu_{ij}, & \text{if } i \neq j. \end{cases} \quad (38)$$

and  $\gamma_{ij}$  is given by (27). Again, this solution can be interpreted in terms of a classical reverse waterfilling scheme.

As in Section III-C, we solve the corresponding dual problem by resorting to the subgradient method. However, now both dual variables  $\lambda_{ij}$  and  $\mu_{ij}$  must be updated as follows

$$\lambda_{ij}^{(t+1)} := \left[ \lambda_{ij}^{(t)} + \alpha_\lambda \left( r_{ij}^{(t+1)} - \sum_{k=j}^{d+i-1} \log(1 + |h_k|^2 p_k^{(t+1)}) \right) \right]^+, \quad (39)$$

<sup>6</sup>As expected, these expressions simplify to the ones for the delay-constrained scenario for  $d = 1$ .

---

**Algorithm 2** Optimal power and rate allocation for the delay-tolerant case.

---

- 1: **Initialize:**  $\{\lambda_{ij}^{(t)}\} := 0$ ,  $\{\mu_{ij}^{(t)}\} := 0$ .
  - 2: **Step 1:** For all  $i$ , allocate power.
  - 3:  $p_i^{(t+1)} := \left[ \frac{\sum_{k=i-d+1}^K \sum_{l=1}^k \lambda_{kl}^{(t)}}{\sum_{j=i}^K \beta_j} - \frac{1}{|h_i|^2} \right]^+$
  - 4: **Step 2:** For all  $i, j$ , cumulative rate allocation.
  - 5:  $r_{ij}^{(t+1)} := \left[ \log \left( \frac{\gamma_{ij}}{\lambda_{ij}^{(t)} + \bar{\mu}_{ij}^{(t)}} \right) \right]^+$
  - 6: **Step 3:** For all  $i, j$ , update multipliers.
  - 7:  $\lambda_{ij}^{(t+1)} := \left[ \lambda_{ij}^{(t)} + \alpha_\lambda \left( r_{ij}^{(t+1)} - \sum_{k=j}^{d+i-1} \log(1 + |h_k|^2 p_k^{(t+1)}) \right) \right]^+$
  - 8:  $\mu_{ij}^{(t+1)} := \left[ \mu_{ij}^{(t)} + \alpha_\mu \left( r_{ij}^{(t+1)} - \sum_{k=j}^i r_{kk}^{(t+1)} \right) \right]^+$
  - 9:  $\mu_{ij}^{(t+1)} := \left[ \mu_{ij}^{(t)} + \alpha_\mu \left( r_{ij}^{(t+1)} - \sum_{k=j}^i r_{kk}^{(t+1)} \right) \right]^+$
  - 10: **Step 4:** Go to Step 1 until stopping criteria is met.
- 

$$\mu_{ij}^{(t+1)} := \left[ \mu_{ij}^{(t)} + \alpha_\mu \left( r_{ij}^{(t+1)} - \sum_{k=j}^i r_{kk}^{(t+1)} \right) \right]^+, \quad (40)$$

where  $\alpha_\lambda$  and  $\alpha_\mu$  denote the corresponding step sizes. Algorithm 2 details the proposed procedure to obtain the optimal power and cumulative rate allocation.

## V. NUMERICAL RESULTS

In this section, we assess the performance of the proposed optimal power and rate allocation schemes. We are particularly interested in analyzing the impact of the correlation coefficient  $\rho$  and the delay  $d$  in the resulting transmission policies. For this reason, in all numerical results we have set the channel gains to a (constant) unit value. Unless otherwise stated, the simulation setup considers a system with  $K = 10$  time slots and an (arbitrary) energy harvesting profile with energy arrivals given by  $E_1 = 0.2$ ,  $E_3 = 0.6$ ,  $E_6 = 0.8$  and  $E_7 = 1.4$ .

### A. Delay-Constrained Scenario ( $d = 1$ )

The resulting optimal power allocation policy is shown in Figure 5. For uncorrelated sources ( $\rho = 0$ ), the optimal policy turns out to be the well-known geometric solution of [3] and [33]. This corresponds to the tightest string below the cumulative energy harvesting curve connecting the original and the total harvested energy by the end of time slot  $K$ . However, as the correlation increases, the harvested energy tends to be spent (i.e., allocated as transmit power) sooner. As a result, in Fig. 5 the slope of the energy consumption curves right after new energy arrivals (e.g., in the beginning of time slot 3) increases with  $\rho$ . This indicates that, in order to minimize the average distortion, one should encode the observations as accurately as possible when some new energy is made available. This stems from the fact that past observations are used here as side information at the receiver. Intuitively, the earlier an observation is accurately encoded,

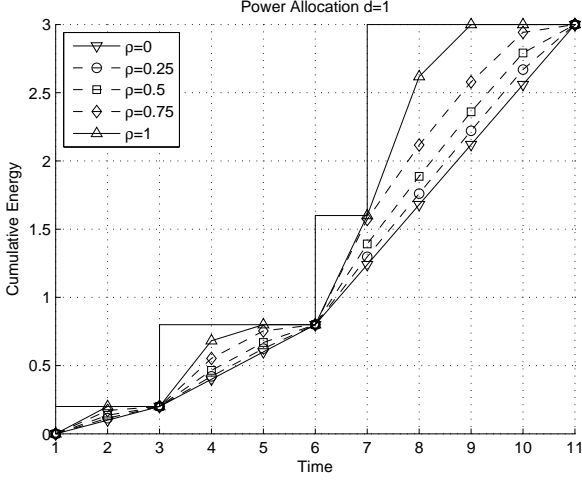


Fig. 5. Optimal power allocation for  $d = 1$  and varying correlation coefficient  $\rho$ .

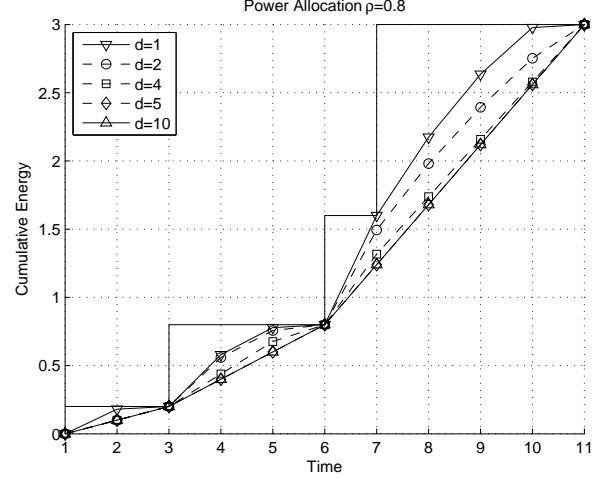


Fig. 7. Optimal power allocation for  $\rho = 0.8$  and varying delay  $d$ .

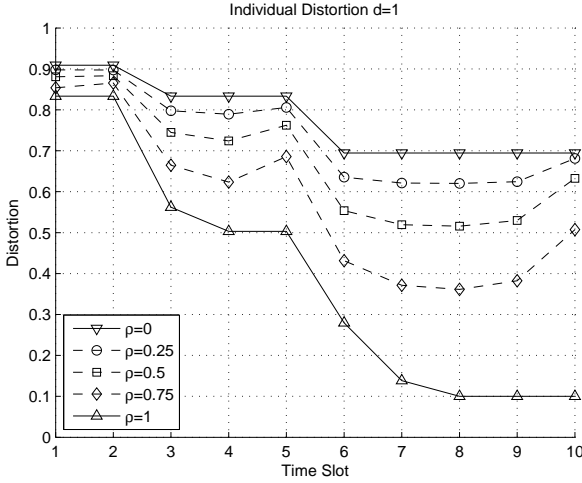


Fig. 6. Individual distortion for  $d = 1$  and varying  $\rho$ .

the more estimates (in subsequent time slots) can benefit from such an increased accuracy. This holds true even at the expense of a reduced (or zero, as in time slot 10, for  $\rho = 1$ ) transmit power being allocated to some subsequent time slots. That is, at the expense of suspending data transmission. All the above is in stark contrast with the uncorrelated case studied in [3] where transmit power is (i) strictly positive for all time slots and (ii) a monotonically increasing function.

Figure 6 depicts the reconstruction distortion for *each* source (and time slot since  $d = 1$ ) associated to the optimal policy. Unsurprisingly, the higher the correlation, the more predictable the sources become and, hence, the lower the distortion (curves are shifted downwards). For correlated sources, however, distortion does not monotonically decrease with time slot index. As discussed in the previous paragraph, this stems from the *anticipated* consumption of the harvested energy for the encoding of previous observations. Consequently, one can observe (i) a substantial decrease of the individual distortion for sources in time slots with energy arrivals (time slots 3,

6, and 7); and (ii) distortion upturns in time slots where the energy harvested so far has been spent or is close to (time slots 2, 5, and 10). Still, the average distortion is lower.

#### B. Delay-Tolerant Scenario ( $d > 1$ )

Figure 7 illustrates the impact of delay on the optimal power allocation. Interestingly, as  $d$  increases the solution converges to the tightest string below the cumulative energy harvesting curve of [3]. The intuition behind is as follows. To recall, the tightest string solution attempts to maximize the total throughput (rate) for the whole transmission period. To that aim, the sequence of transmit powers (and rates) must be monotonically increasing, that is, transmit power is higher by the end of the transmission period (i.e., last time slot(s)). For  $d = 1$ , on the contrary, the allocated transmit power (and, thus, rate) is higher in time slots with energy arrivals, and not necessarily in the last one(s). Moreover, the source must be reconstructed *immediately*, that is, after  $d = 1$  time slots (assuming the processing time at the FC to be negligible). In other words, there is some *urgency* to allocate power (namely, spend energy). This is in stark contrast with the tightest string solution where higher power and rates can be found at the *end*. Things, however, are radically different when  $d$  increases. On the one hand, the deadline by which individual sources must be reconstructed is shifted  $d$  time slots towards the *end*. On the other, the rates (and power) needed to encode a specific source can be allocated over *multiple* time slots, rather than just one. Hence, for increasing  $d$  the urgency to allocate power decreases and, thus, the way in which power is allocated is more aligned with that of the tightest string solution.

Figure 8 depicts the average distortion as a function of delay. Clearly, the average distortion decreases with delay since the higher the delay, the higher the degrees of freedom to allocate transmit power (and, thus, spend energy in a more sensible manner). Unsurprisingly, distortion is lower for higher values of  $\rho$ , since the preceding (correlated) sources used as side information at the FC are more informative.

Next, we investigate to what extent our system leverages on



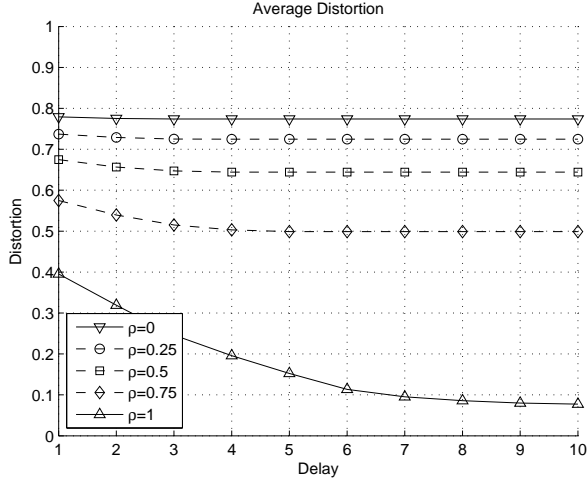


Fig. 8. Average distortion vs. delay  $d$  for varying correlation  $\rho$ .

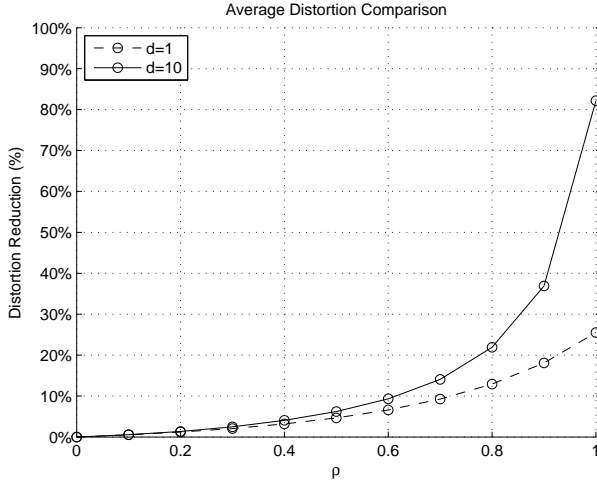


Fig. 9. Reduction in average distortion.

the knowledge on source correlation. To that aim, the rate and power allocation policy from [16], which was derived for a scenario with *uncorrelated* sources, is used as a benchmark. Specifically, whereas the source encoding rate depends on  $I(x_i; u_i)$  (namely, the mutual information with the *current* source only) the sources at the FC are reconstructed according to (9). Our approach, on the contrary, exploits correlation both in the encoding and decoding/reconstruction processes. Figure 9 shows the normalized reduction (difference) in the average distortion attained by such benchmark and our scheme. For delay-constrained scenarios, the reduction in distortion can be as high 25% for our scheme. For delay-tolerant ones ( $d = 10$ ), reduction can go up to 80%, which is very remarkable.

### C. Comparison with an Online Policy

As discussed earlier, the proposed (offline) transmit power and rate allocation scheme requires *non-causal* knowledge on energy arrivals. Here, instead, we introduce a more realistic *online* version just requiring *causal* knowledge. The offline scheme will be used as a benchmark.

Similar to [6], a *myopic*<sup>7</sup> online policy can be computed as follows. Assume for a moment that, after harvesting some energy in the initial timeslot (i.e.,  $E_1 > 0$ ), no additional energy is harvested in subsequent timeslots. Hence, we let  $E_2 = \dots = E_K = 0$  and solve problem (34) for  $k = 1, \dots, K$ . In the absence of knowledge on future energy arrivals, this is a sensible approach too. After all, distortion would be minimized should no additional energy be actually harvested. And, otherwise, we can react accordingly. Let  $k_0 < K$  denote the next timeslot in which some energy is harvested (i.e.,  $E_{k_0} > 0$ ). For the preceding timeslots (i.e.,  $k = 1, \dots, k_0 - 1$ ), we force the power and rate allocations computed after the last energy arrival to remain unchanged. Hence, the unspent energy in the beginning of timeslot  $k_0$  reads  $E_{k_0}^u = \sum_{j=1}^{k_0-1} E_j - \sum_{j=1}^{k_0-1} p_j$ . Next, we let  $E_{k_0} := E_{k_0}^u + E_{k_0}$  and  $E_{k_0+1} = \dots = E_K = 0$  and, again, solve problem (34) for  $k = k_0, \dots, K$ . That is, we compute the optimal power and rate allocations for all subsequent timeslots. This procedure is iterated until all energy arrivals have been accounted for.

Of course, no optimality can be claimed for the resulting policy. Still, the interesting property of such scheme is its ability to adjust (re-compute) the remaining power and rate allocations every time that some energy is harvested. By doing so, the additional (and causal) knowledge on energy arrivals is effectively exploited.

Figure 10 illustrates the performance of the offline and online policies vs. the intensity rate of energy arrivals (which are modeled as a Poisson process). Unsurprisingly, the distortion of the offline versions turns out to be a lower bound of that attained by online ones. For a given intensity rate, the additional distortion associated to the online version can be regarded as moderate (some 20% at an energy arrival rate equal to 1,  $\rho = 0.2$ , and  $d = 1$ ). Interestingly, the online version requires a 40% increase of the intensity rate to achieve the same distortion as its offline counterpart (for the same operating point). The distortion gap becomes narrower for delay-tolerant scenarios ( $d = 10$ ) and wider in percentage for scenarios with high correlation (see  $\rho = 0.8$  curves) or when the intensity rate increases.

### D. Convergence

Next, we investigate the convergence properties of the proposed scheme. Specifically, in Figure 11 we depict the relative error  $\varepsilon$  between the average distortion at iteration  $t$  and its optimal value, namely,  $\varepsilon = |D_{avg}^* - D_{avg}^{(t)}| / D_{avg}^*$ . For the update of the dual variables in (31), we have used a time-varying step size<sup>8</sup>. Clearly, convergence is slower for larger  $d$  values. This stems from the fact that, for delay-tolerant scenarios, the search space for the solution is larger, as the summation in equation (34c) evidences.

<sup>7</sup>More general online policies accounting for different degrees of availability of channel and energy state information can be also considered (see e.g., [34]).

<sup>8</sup>The step size used is  $\alpha^{(t)} = \bar{\alpha}^{(t)} / \|g^{(t)}\|_2$ , where  $g^{(t)}$  is the corresponding subgradient and  $\bar{\alpha} = 1/\sqrt{t}$ . This diminishing step size satisfies the convergence conditions given by  $\bar{\alpha}^{(t)} \geq 0$ ,  $\lim_{t \rightarrow \infty} \bar{\alpha}^{(t)} = 0$  and  $\sum_{t=1}^{\infty} \bar{\alpha}^{(t)} = \infty$  [22].

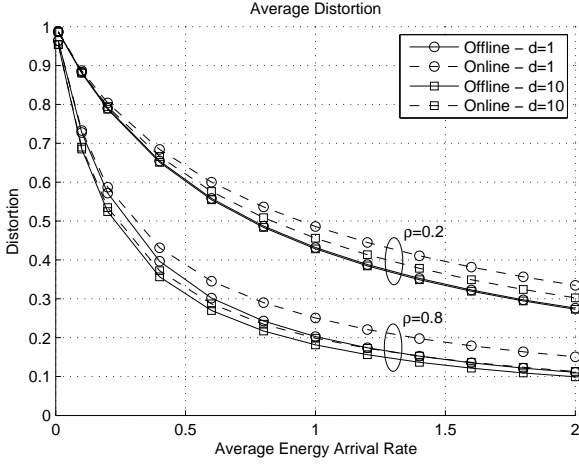


Fig. 10. Average distortion for the offline and online policies, for low ( $\rho = 0.2$ ) and high ( $\rho = 0.8$ ) correlation, and delay-constrained ( $d = 1$ ) and delay-tolerant ( $d = 10$ ) scenarios.

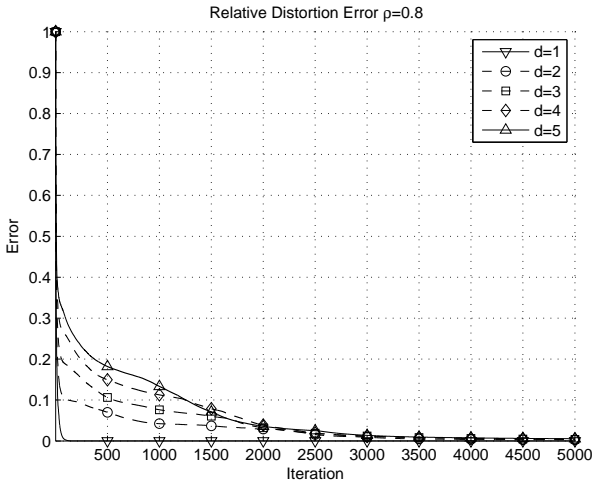


Fig. 11. Relative distortion error for  $\rho = 0.8$  and varying  $d$ .

### E. Allocation of Individual Rates

As discussed in Section IV, for  $d > 1$  there exists an infinite number of solutions for the allocation of the *individual* rates (the system of equations is underdetermined). In order to get some insight on how individual rates are allocated, we will select the solution with the lowest 2-norm. This, clearly, penalizes solutions with very large (dissimilar) rates.

After solving the optimization problem (34) and determining the optimal cumulative rates  $r_{i,j}$ , we find the individual

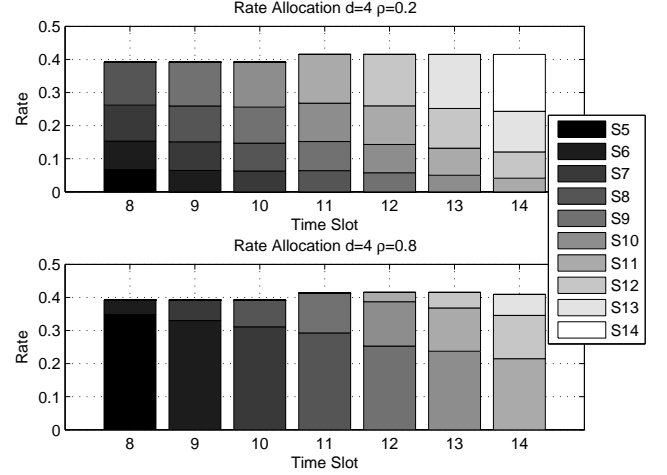


Fig. 12. Allocation of individual rates for sources with low (top) and high (bottom) correlation ( $K = 20$ ;  $d = 4$ ; energy profile:  $E_1 = 0.2$ ,  $E_2 = 1$ ,  $E_4 = 0.6$ ,  $E_6 = 1$ ,  $E_7 = 0.8$ ,  $E_8 = 0.2$ ,  $E_9 = 0.4$ ,  $E_{11} = 1.4$ ,  $E_{13} = 0.6$ ,  $E_{14} = 0.6$ ,  $E_{16} = 0.8$ ,  $E_{17} = 0.2$ ,  $E_{18} = 1$ ,  $E_{19} = 0.2$ , and  $E_{20} = 0.4$ ).

rates  $R_{i,j}$  by solving:

$$\min_{\{R_{i,j}\}} \left( \sum_{i=1}^K \sum_{j=i-d+1}^i R_{i,j}^2 \right)^{1/2} \quad (41a)$$

$$\text{s.t. } r_{ii} = \sum_{j=i}^{i+d-1} R_{j,i}, i = 1, \dots, K \quad (41b)$$

$$\sum_{j=i-d+1}^i R_{i,j} \leq \log(1 + |h_i|^2 p_i), i = 1, \dots, K \quad (41c)$$

$$R_{i,j} \geq 0, i = 1, \dots, K, j = i - d + 1, \dots, i \quad (41d)$$

To that aim, we need to use  $r_{ii}$  and  $p_i$  from the solution of the (cumulative) rate and power allocation problem as an input (see first and second inequality constraints in the problem above).

Figure 12 shows the allocation of *individual* rates over the  $d$  time slots for each source (a different color is used for each source). We consider scenarios with sources exhibiting low ( $\rho = 0.2$ ) and high ( $\rho = 0.8$ ) correlation. Interestingly enough, the higher the correlation, the lower the spread of individual rates over time slots (fewer sources in each time slot). This is consistent with the fact that, as discussed earlier, for low  $\rho$  (and  $d = 1$ ) energy tends to be spent sooner. Accordingly, in delay-tolerant scenarios where the encoded data is transmitted in a number of time slots, when correlation is high the first time slots are favored.

## VI. CONCLUSIONS

In this paper, we have investigated the impact of source correlation in the design of point-to-point optimal transmission policies with energy-harvesting sensors. We have considered both delay-constrained delay-tolerant scenarios. In both cases, our goal was to minimize the average distortion in the decoded

(reconstructed) observations by using data from previously encoded sources as side information. We have formulated the problems in a convex optimization framework. Besides, we have proposed an iterative procedure, based on the subgradient method, to solve both problems. Interestingly, the procedure entails the interaction of a directional and reverse water-filling schemes in each iteration. For the delay-constrained scenario, numerical results revealed that, differently from the uncorrelated case, minimizing the average distortion implies encoding observations as accurately as possible upon energy arrivals. This holds true even if the transmit power allocated to subsequent time slots is lower or, eventually, zero (and, thus, an increase in distortion in such time slots). For the delay-tolerant scenario, we have observed that as delay increases, the power allocation policy converges to the tightest string below the cumulative energy harvesting curve. And, also, that the average distortion decreases. In comparison with other schemes not exploiting correlated sources as side information, ours attains an average distortion which is substantially lower (with reductions of up to 25% or 80% for  $d = 1$  and  $d = 10$ , respectively). We have also proposed a myopic online policy exhibiting a moderate performance gap (some 20% for low correlation and delay-constrained scenarios) with respect to the offline (optimal) policy. Besides, we have found that the time needed for the algorithm to convergence is higher for delay-tolerant scenarios since the search space is substantially larger there. Finally, we have observed that for delay-tolerant scenarios, the higher the correlation, the lower the spread of individual rates over time slots.

#### APPENDIX A

##### DERIVATION OF THE AVERAGE DISTORTION IN (11)

For compactness, hereinafter we let  $R_i \triangleq \sum_{j=i}^{i+d-1} R_{j,i}$  denote the rate assigned to the  $i$ -th source over its  $d$  time slots; and  $D_i = \sigma_{x_i|u_1, \dots, u_i}^2$  the distortion for the  $i$ -th source which can be recursively expressed as [35]:

$$\sigma_{x_k|u_1, \dots, u_k}^2 = \sigma_{x_k|u_1, \dots, u_{k-1}}^2 - \frac{\text{cov}^2(u_k, x_k|u_1, \dots, u_{k-1})}{\sigma_{u_k|u_1, \dots, u_{k-1}}^2}.$$

We prove by induction that

$$\begin{aligned} D_i &= \sigma_x^2 \left( (1 - \rho) \sum_{j=2}^i \rho^{i-j} e^{-\sum_{k=j}^i \sum_{l=k}^{k+d-1} R_{l,k}} \right. \\ &\quad \left. + \rho^{i-1} e^{-\sum_{k=1}^i \sum_{l=k}^{k+d-1} R_{l,k}} \right) \\ &= \sigma_x^2 \left( (1 - \rho) \sum_{j=2}^i \rho^{i-j} e^{-\sum_{k=j}^i R_k} \right. \\ &\quad \left. + \rho^{i-1} e^{-\sum_{k=1}^i R_k} \right). \end{aligned} \quad (42)$$

We start by showing this expression holds for the base case ( $i = 1$ ). That is

$$\begin{aligned} D_1 &= \sigma_{x_1|u_1}^2 \\ &= \sigma_x^2 - \frac{(\sigma_x^2)^2}{\sigma_x^2 + \sigma_{z_1}^2} = \sigma_x^2 - \frac{(\sigma_x^2)^2}{\sigma_x^2 + \frac{\sigma_x^2}{e^{R_1} - 1}} = \sigma_x^2 e^{-R_1}, \end{aligned}$$

which satisfies expression (42). For the inductive step, assume expression (42) is true for  $i = n$ . Then consider

$$\begin{aligned} D_{n+1} &= \sigma_{x_{n+1}|u_1, \dots, u_{n+1}}^2 \\ &= \sigma_{x_{n+1}|u_1, \dots, u_n}^2 - \sigma_{x_{n+1}|u_1, \dots, u_n}^2 (1 - e^{-R_{n+1}}) \\ &= \sigma_{x_{n+1}|u_1, \dots, u_n}^2 e^{-R_{n+1}} \\ &= \sigma_{\sqrt{\rho}x_n + w_n|u_1, \dots, u_n}^2 e^{-R_{n+1}} \\ &= \left( \rho \sigma_{x_n|u_1, \dots, u_n}^2 + \sigma_{w_n|u_1, \dots, u_n}^2 \right) e^{-R_{n+1}} \\ &= (\rho D_n + \sigma_x^2 (1 - \rho)) e^{-R_{n+1}} \\ &= D_n \rho e^{-R_{n+1}} + \sigma_x^2 (1 - \rho) e^{-R_{n+1}} \end{aligned}$$

Then by the induction hypothesis we have

$$\begin{aligned} D_{n+1} &= \sigma_x^2 \left( (1 - \rho) \sum_{j=2}^n \rho^{n-j} e^{-\sum_{k=j}^n R_k} \right. \\ &\quad \left. + \rho^{n-1} e^{-\sum_{k=1}^n R_k} \right) \rho e^{-R_{n+1}} + \sigma_x^2 (1 - \rho) e^{-R_{n+1}} \end{aligned}$$

and by rearranging terms we have

$$\begin{aligned} D_{n+1} &= \sigma_x^2 \left( (1 - \rho) \sum_{j=2}^{n+1} \rho^{n+1-j} e^{-\sum_{k=j}^{n+1} R_k} \right. \\ &\quad \left. + \rho^{n+1-1} e^{-\sum_{k=1}^{n+1} R_k} \right). \end{aligned}$$

Thus, expression (42) holds for  $i = n + 1$ . Therefore, by the principle of induction, expression (42) holds for all  $i$ .

#### APPENDIX B

##### PROOF OF PROPOSITION 2.

*Proof:* For notational convenience, let  $c_i \triangleq \log(1 + |h_i|^2 p_i)$  denote the channel capacity in the  $i$ -th timeslot. First, we focus on the direct proof. Assuming that the LHS of (30), which is given by the system of inequalities

$$\sum_{j=i}^{i+d-1} R_{j,i} = r_{ii}, \quad i = 1, \dots, K, \quad (43a)$$

$$\sum_{j=i-d+1}^i R_{i,j} \leq c_i, \quad i = 1, \dots, K, \quad (43b)$$

$$R_{i,j} \geq 0, \quad i = 1, \dots, K, j = 1, \dots, i - d + 1, \quad (43c)$$

has a solution in terms of individual rates  $R_{i,j}$ , our goal is to find the system of inequalities in the RHS of (30), namely

$$r_{ij} = \sum_{k=j}^i r_{kk}, \quad i = 1, \dots, K, j = 1, \dots, i - 1, \quad (44a)$$

$$r_{ij} \leq \sum_{k=j}^{i+d-1} c_k, \quad i = 1, \dots, K, j = 1, \dots, i, \quad (44b)$$

$$r_{ij} \geq 0, \quad i = 1, \dots, K, j = 1, \dots, i. \quad (44c)$$

The constraints (44a) follow directly from the definition of the cumulative rates (33). Constraint (44c) is also straightforward since, from its definition in (32), the cumulative rates  $r_{i,j}$  can be expressed as a summation of *non-negative* (see (43c)) individual rates  $R_{i,j}$ . As for (44b), note that from (43b) each *non-negative* individual rate can be upper-bounded as follows

$$R_{i,j} \leq c_i, \quad i = 1, \dots, K, j = 1, \dots, i - d + 1. \quad (45)$$

Next, by direct substitution of the bounds (45) into the equalities (43a), we have that

$$r_{ii} \leq \sum_{j=i}^{i+d-1} c_j, \quad i = 1, \dots, K \quad (46)$$

Finally, by substitution of (46) into the definition of cumulative rates (33), inequality (44b) follows.

Consider now the converse. Assume that the RHS of (30), which is also given by the system of inequalities (44), has a solution in terms of cumulative rates  $r_{ij}$ . Then, we want to prove that there exists a non-empty set of individual rates  $R_{i,j}$  satisfying the inequalities (43) (i.e., the RHS of (30) has a solution even if it might not be unique, as discussed earlier). To prove that, we focus on the more restrictive case where we force the capacity constraint (43b) to be satisfied with equality. Hence, the first two constraints in (44) become:

$$\sum_{j=i}^{i+d-1} R_{j,i} = r_{ii}, \quad i = 1, \dots, K \quad (47a)$$

$$\sum_{j=i-d+1}^i R_{i,j} = c_i, \quad i = 1, \dots, K. \quad (47b)$$

The system of equations above can be rewritten in matrix form:

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (48)$$

where we have defined the column vectors  $\mathbf{x} \triangleq [R_{1,1}, R_{2,1}, \dots, R_{K,K}]^T$  and  $\mathbf{b} \triangleq [r_{1,1}, \dots, r_{K,K}, c_1, \dots, c_K]^T$ , and where matrix  $\mathbf{A}$  is given by the  $\{0, 1\}$  entries yielding the summations in (47). Next, we resort to Farkas' lemma:

**Lemma 3** (Farkas' Lemma [21]). *If  $\mathbf{A} \in \mathbb{R}^{m \times n}$  and  $\mathbf{b} \in \mathbb{R}^m$ , then exactly one of the following holds:*

- (i) *There exists  $\mathbf{x} \in \mathbb{R}^n$  such that  $\mathbf{A}\mathbf{x} = \mathbf{b}$  and  $\mathbf{x} \geq 0$ .*
- (ii) *There exists  $\mathbf{y} \in \mathbb{R}^m$  such that  $\mathbf{y}^T \mathbf{A} \geq 0$  and  $\mathbf{y}^T \mathbf{b} < 0$ .*

where the inequality  $\mathbf{x} \geq 0$  is defined element-wise. Clearly, alternative (i) in the Farkas lemma states that, if it holds, a solution to the LHS in terms of individual rates exists. In the next paragraphs, we prove (by contradiction) that alternative (ii) does *not* hold for our problem. To that aim, we start by defining  $\mathbf{y} \triangleq [k_{r_{11}}, \dots, k_{r_{KK}}, k_{c_1}, \dots, k_{c_K}]^T$ . Assume that alternative (ii) holds. To satisfy the condition  $\mathbf{y}^T \mathbf{A} \geq 0$ , there must exist a nonnegative set of coefficients  $k_{r_{ii}}$  and  $k_{c_i}$  such that

$$k_{r_{ii}} + k_{c_j} \geq 0, \quad i = 1, \dots, K, j = i, \dots, i + d - 1. \quad (49)$$

And, condition  $\mathbf{y}^T \mathbf{b} < 0$  can be rewritten as

$$\sum_{i=1}^K k_{r_{ii}} r_{ii} + \sum_{i=1}^K k_{c_i} c_i < 0. \quad (50)$$

Next, we will check that for any set of valid  $k_{r_{ii}}$  and  $k_{c_j}$  equation (50) does not hold. To that aim, we will determine the lowest possible value of the LHS of (50) subject to the inequalities given by (49). In other words, we need to solve an optimization (minimization) problem with the LHS of (50) playing the role of the objective function and (49) as constraints. Since  $r_{ii}$  and  $c_i$  are nonnegative, this is a linear program. Hence, the solution will lie at the vertex of the feasible region defined by (49) [32, Chapter 7]. Since the

expressions (49) define a convex cone, its only vertex is given by

$$k_{r_{ii}} + k_{c_j} = 0, \quad i = 1, \dots, K, j = i, \dots, i + d - 1. \quad (51)$$

By recursively analyzing the various equations in (51), we conclude that necessarily

$$k \triangleq k_{r_{ii}} = -k_{c_i}, \quad i = 1, \dots, K. \quad (52)$$

That is, except for the sign, all the coefficients are identical. By replacing (52) into the LHS of (50), the objective function in the optimization problem becomes

$$k \sum_{i=1}^K r_{ii} - k \sum_{i=1}^K c_i. \quad (53)$$

From (44a) and (44b), we have that  $\sum_{i=1}^K r_{ii} \leq \sum_{i=1}^K c_i$ . That is, the sum of cumulative rates of all sources is *below or equal to* the channel capacity over all time slots. However, since the objective function in (34) is nonincreasing in all  $r_{ij}$ , the optimal solution of (34) must satisfy  $\sum_{i=1}^K r_{ii} = \sum_{i=1}^K c_i$  (i.e., with equality). This means that, necessarily, (53) is lower bounded by 0, hence, (50) does not hold and, in turn, alternative (ii) in the Farkas theorem does not hold either. This concludes the proof. ■

## REFERENCES

- [1] R. J. Vullers, R. Schaijk, H. J. Visser, J. Penders, and C. V. Hoof, "Energy harvesting for autonomous wireless sensor networks," *IEEE Solid-State Circuits Magazine*, vol. 2, no. 2, pp. 29–38, 2010.
- [2] S. Ulukus, A. Yener, E. Erkip, O. Simeone, M. Zorzi, P. Grover, and K. Huang, "Energy harvesting wireless communications: A review of recent advances," *IEEE Journal on Selected Areas in Communications*, vol. PP, no. 99, pp. 1–1, 2015.
- [3] J. Yang and S. Ulukus, "Optimal packet scheduling in an energy harvesting communication system," *IEEE Transactions on Communications*, vol. 60, no. 1, pp. 220–230, 2012.
- [4] K. Tutuncuoglu and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 1180–1189, 2012.
- [5] B. Devillers and D. Gündüz, "A general framework for the optimization of energy harvesting communication systems with battery imperfections," *IEEE Journal of Communications and Networks*, vol. 14, no. 2, pp. 130–139, 2012.
- [6] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: Optimal policies," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1732–1743, 2011.
- [7] O. Orhan, D. Gündüz, and E. Erkip, "Optimal packet scheduling for an energy harvesting transmitter with processing cost," in *IEEE International Conference on Communications (ICC)*, Budapest, Hungary, June 2013, pp. 3110–3114.
- [8] O. Ozel, J. Yang, and S. Ulukus, "Optimal broadcast scheduling for an energy harvesting rechargeable transmitter with a finite capacity battery," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2193–2203, 2012.
- [9] J. Yang and S. Ulukus, "Optimal packet scheduling in a multiple access channel with energy harvesting transmitters," *Communications and Networks, Journal of*, vol. 14, no. 2, pp. 140–150, 2012.
- [10] S. Cui, J.-J. Xiao, A. J. Goldsmith, Z.-Q. Luo, and H. V. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4683–4695, 2007.
- [11] S. C. Draper and G. W. Wornell, "Side information aware coding strategies for sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 966–976, 2004.
- [12] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 1–10, 1976.
- [13] P. Castiglione, O. Simeone, E. Erkip, and T. Zemen, "Energy management policies for energy-neutral source-channel coding," *Communications, IEEE Transactions on*, vol. 60, no. 9, pp. 2668–2678, 2012.

- [14] P. Castiglione and G. Matz, "Energy-neutral source-channel coding with battery and memory size constraints," *IEEE Transactions on Communications*, vol. 62, no. 4, pp. 1373–1381, April 2014.
- [15] C. Tapparello, O. Simeone, and M. Rossi, "Dynamic compression-transmission for energy-harvesting multihop networks with correlated sources," *IEEE/ACM Transactions on Networking*, vol. 22, no. 6, pp. 1729–1741, Dec 2012.
- [16] O. Orhan, D. Gündüz, and E. Erkip, "Delay-constrained distortion minimization for energy harvesting transmission over a fading channel," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, Istanbul, Turkey, July 2013, pp. 1794–1798.
- [17] R. Puri, A. Majumdar, P. Ishwar, and K. Ramchandran, "Distributed video coding in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 94–106, 2006.
- [18] P. Ishwar and K. Ramchandran, "On decoder-latency versus performance tradeoffs in differential predictive coding," in *IEEE International Conference on Image Processing (ICIP)*, Singapore, October 2004, pp. 1097–1100.
- [19] N. Ma and P. Ishwar, "On delayed sequential coding of correlated sources," *IEEE Transactions on Information Theory*, vol. 57, no. 6, pp. 3763–3782, 2011.
- [20] E.-H. Yang, L. Zheng, D.-K. He, and Z. Zhang, "Rate distortion theory for causal video coding: Characterization, computation algorithm, and comparison," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5258–5280, 2011.
- [21] M. Calvo-Fullana, J. Matamoros, and C. Antón-Haro, "Reconstruction of correlated sources with energy harvesting constraints," in *European Wireless (EW)*, Budapest, May 2015, pp. 1–6.
- [22] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.
- [23] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2012.
- [24] J. Xu and R. Zhang, "Throughput optimal policies for energy harvesting wireless transmitters with non-ideal circuit power," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 2, pp. 322–332, 2014.
- [25] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2349–2360, 2005.
- [26] Z. Wei, X. Zhu, S. Sun, Y. Huang, L. Dong, and Y. Jiang, "Full-duplex versus half-duplex amplify-and-forward relaying: Which is more energy efficient in 60-ghz dual-hop indoor wireless systems?" *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 12, pp. 2936–2947, 2015.
- [27] P. Ishwar, R. Puri, K. Ramchandran, and S. S. Pradhan, "On rate-constrained distributed estimation in unreliable sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 765–775, 2005.
- [28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.
- [29] J. Gong, S. Zhou, and Z. Niu, "Optimal power allocation for energy harvesting and power grid coexisting wireless communication systems," *IEEE Transactions on Communications*, vol. 61, no. 7, pp. 3040–3049, 2013.
- [30] C. Hu, J. Gong, X. Wang, S. Zhou, and Z. Niu, "Optimal green energy utilization in mimo systems with hybrid energy supplies," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 8, pp. 3675–3688, 2015.
- [31] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex analysis and optimization*. Athena Scientific, 2003.
- [32] A. Schrijver, *Theory of Linear and Integer Programming*. John Wiley & Sons, 1998.
- [33] M. A. Zafer and E. Modiano, "A calculus approach to energy-efficient data transmission with quality-of-service constraints," *IEEE/ACM Transactions on Networking*, vol. 17, no. 3, pp. 898–911, 2009.
- [34] H. Li, J. Xu, R. Zhang, and S. Cui, "A general utility optimization framework for energy-harvesting-based wireless communications," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 79–85, 2015.
- [35] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall, 1993.